Tampere University of Applied Sciences

Guiding the player in a non-linear environment using sound

Juuso Tolonen

BACHELOR'S THESIS
April 2020

Media and Arts
Music Production

ABSTRACT

Tampereen ammattikorkeakoulu
Tampere University of Applied Sciences
Degree Programme in Media and Arts
Music Production

TOLONEN, JUUSO:
Guiding the Player in a Non-linear Environment Using Sound

Bachelor's thesis 49 pages, appendices 5 pages
April 2020

Through history of videogames to modern day examples, the importance of sound in guiding the player through non-linear environments is apparent. Modern games can have hundreds of hours of gameplay and vast, massive areas. As the player has control of the events and progression of the gameplay, it raises unique challenges only present in videogame development such as player navigation.

In the thesis these challenges were explored through an auditory point of view. The techniques behind them, their implementation and their functionality were studied from various examples and through interviews. The role of audio and its functionality in guiding the player was examined as a part of the game development process.

In the practical project these theories were applied to an example game scene showcasing their functionality and importance in the game world. This also shows what goes into the technical side of implementing these sounds and why that is important for the player experience.

Key words: sound design, game design, video games, game audio

CONTENTS

GLOSSARY

Diegetic sound	A sound that is part of the story world, like a footstep.
Earcon	Short, recognizable sound tied to a certain event for example an error sound.
FPS	First-person Shooter
HPF	Highpass-filter, audio processing where lower frequencies are being filtered
LPF	Lowpass-filter, audio processing where higher frequencies are being filtered
Non-diegetic sound	A sound outside of the story world, for example score music.
NPC	Non-playable character in a videogame.
RTPC	Real-time Parameter Controls, used to control user-defined parameters in Wwise
SFX	Sound effect
VO	Voice-over

1 INTRODUCTION

Sound is an integral part of our navigation in our day-to-day life and as video-games are becoming more and more immersive and realistic, these guiding elements of audio are to be taken into consideration in the non-linear virtual world as well (Jacobsen 2015). It is common for films to use sound to tell a story beyond the screen or call special attention to something but what happens when you don't control the viewers focus? In films the director can guide the viewers and their attention as they please but in videogames the player has control of the events.

As the players are mostly free to do whatever they wish in a game environment it is imperative for the games progression to guide the player subtly towards their next goal. With this uncertainty of the player's decisions, there are unique challenges to guiding the player in the game world. With these challenges the role of audio is emphasized as there is only so much visual information that can be given or received at once. Sound helps to convey information and can be anything from a subtle cue piquing the players interest to a heart racing sign of danger. Sound has been shown to being capable of invoking a significant physical reaction in the player (Usher 2012).

Videogames have grown to be massive open environments with tons of visual information (Chikhani 2015). It is very easy for the player to feel lost and confused in a setting like this without proper auditory feedback and the sense of presence in the game is diminished (Jorgensen 2008). Modern systems also have the resources to give sound to virtually everything in a game environment and so managing and mixing these assets is a skill that is very different from linear media.

PICTURE 1. Combat situation in Control (Remedy Entertainment 2019).

As an example, looking at Picture 1 above it could be easy to miss the enemy on the left advancing on the player. This could be brought to the players attention with sound cues as that enemy is likely the most immediate threat that the player should focus on. In my interview with the lead sound designer at Remedy, Ville Sorsa, he elaborated on the systems behind the audio in Control, which will be used as an extensive example in the thesis (Appendix 2). Using adaptive mixing techniques, the enemies further away could be mixed down or completely culled just to shift more focus to the most immediate threat which would be the closest enemy aiming at you (Sorsa 2020).

2 GAME AUDIO HISTORY

When the first videogames were created they had no sound, due to technical hardware limitations (Collins 2008, 9). When they eventually did include sound in games it was incredibly simplistic due to these same limitations but even the primal sound offered important feedback to the player. One of the first examples would be Computer Space (Nuttin' Associates' 1971) that very ambitiously advertised the game's capabilites with a flier that said: *"The thrust motors from your rocket ship, the rocket turning signals, the firing on of your missiles and explosions fill the air with the sights and sounds of combat as you battle against the saucers for the highest score."* (Collins 2008, 8).

2.1 The role of audio in video games

Creating audio for video games is a unique challenge in the sense that nothing in a game world is real, nothing is emitting sound unless you tell it to. Every single detail in the world must be acknowledged and designed by the sound designer. (Füsslin 2017). This is essential to make the player feel that the world is truly responsive.

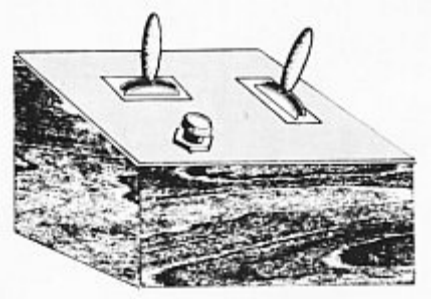
Due to the massive leaps in technology in the past decade, videogames have become incredibly immersive storytellers. Sound is an integral part of this immersion, giving the player auditory feedback of the world around them (Grimshaw 2008). Sounds can convey information, but they can also tell a story. For example, a simple footstep can tell you the weight of your characters armour giving the player hints about how strong their current defences are. If you then add an encumbered whimper on every footstep, it tells a whole different story of the character itself.

Sounds are also a strong way to reinforce player actions (Adams 2006). A positive cadence will instantly reward and inform the player at the same time. On the other hand, a jarring, dissonant or an unpleasant sound is telling the player to avoid doing the same action again. Sounds can either further excite the player or calm the situation down. This also brings up the notion of aesthetic choice. When designing sounds, it is always a careful balance between aesthetics and realism.

Sometimes it is important to make an overemphasis on the aesthetic side just to make the game more enjoyable and give more rewarding feedback to the player. If a gun sounds nice, you want to shoot it more.

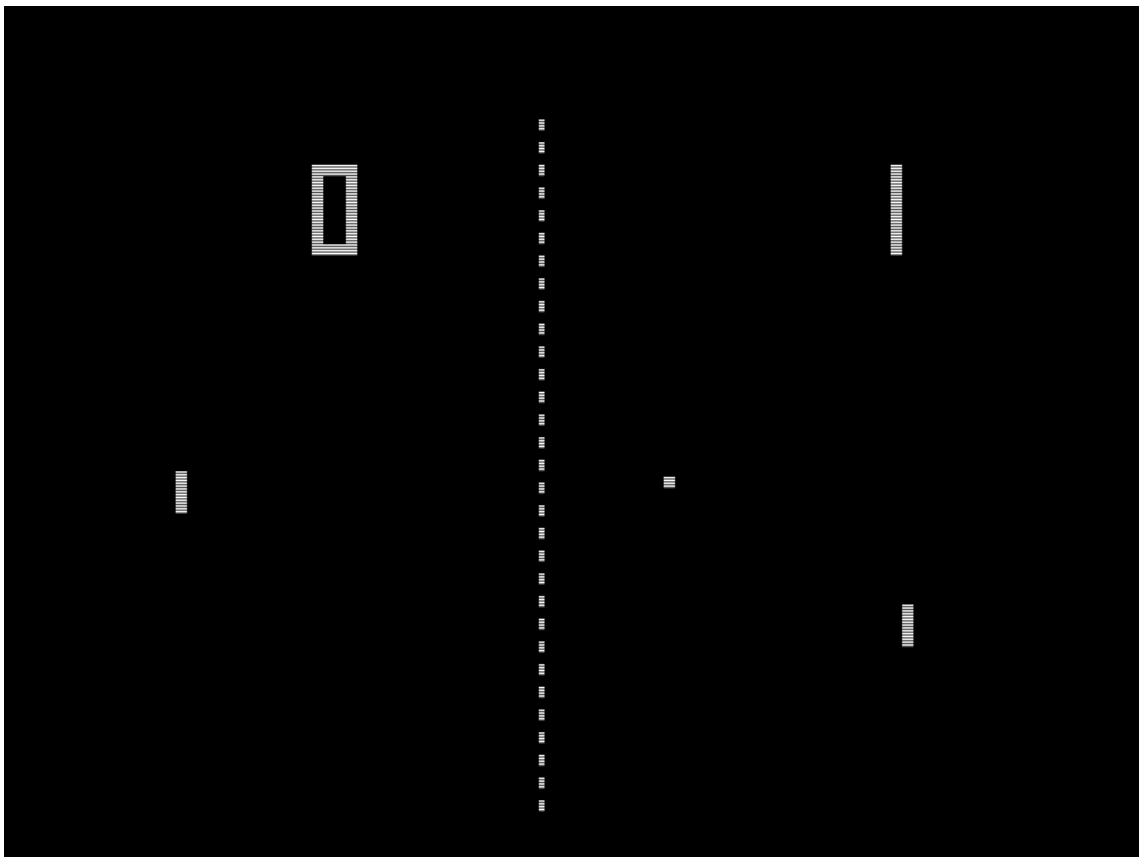
2.1.1 In the past

Spacewar! (Steve Russell 1962) is considered one of the first real-time video-games and while the initial version of the game did not itself have sounds there is an interesting story relating to the controllers created for the game. Joysticks did not exist back then, so Alan Koton and Robert A. Saunders designed a custom control box for the game. One of the design features was that the button for firing torpedoes had to be silent so that the other player wouldn't hear when you were trying to fire. (Graetz 1981, Creative Computing.)



PICTURE 2. Sketch of the original design for Spacewar! control box. (Graetz 1981).

Pong (Atari 1972) was one the first games to utilize sound and even with its very crude audio capabilities it provided very important feedback to the player (Collins 2008, 8). By varying the pitch and length of a simple beep the game could inform the player whether the ball hit a wall, the paddle or if a point was scored.



PICTURE 3. Pong (Atari, 1972).

A real-life example of guiding with audio would be in the early arcades where sound was used to literally guide the player to the machine (Collins 2008, 9). The machines themselves emitted noise with flashy lights to attract customers to the machines. The games themselves still had very rudimentary sound capabilities but the mention of sound effects was used as a selling feature (Collins 2008, 7). After the success of Pong, more ambitious videogame projects emerged but due to technological limitations these early videogames could be sometimes very hard to read visually. Audio played a crucial part of giving the player more feedback for example when taking damage or collecting score points.

2.1.2 Sound in games compared to linear media

In linear media such as films the story arc and pacing of the viewers experience has been determined and is controlled completely by the director. On every viewing the experience is the same. You know exactly when a certain event happens,

and every sound in that certain event has been carefully laid out and is going to play the same way every time.

Videogames give the control to the player thus immersing them further into the world. You as the player decide when and how certain events are going to play out and you control your own progression in the game world. This means you as the developer cannot necessarily control when a certain sound happens in the game, for example a jumping sound. If the player keeps repeating the same action, in this case the jump, it would be like watching the same short scene from a movie repeatedly. This might become annoying and break you out of the immersion completely if everything is always exactly the same (Jacobsen 2018). To combat this listening fatigue, you would add variations to the sound. You could add a random vocalization that play occasionally, and the actual jumping sound would be created by having multiple layers of sound assets each with different variations.

Bjorn Jacobsen, a sound designer who has worked on multiple major game releases talks about the nature of communication in his talk “Video game and real life audio as game design” (2015). Sound can be a tool of communication. One form of communication could be described as monologue which would be a continuous source of information from the source to the listener. There are also dialogue where all the parties involved exchange information, consultation where one source is gathering information from different sources and observation where you pick up information passively. With these terms in mind, films could be considered monologue, but in videogames you utilize all of these communication tools. (Jacobsen 2015).

Diegetic and non-diegetic are terms often used in film theory for sounds accompanying visuals. Diegetic sound would have a visible source on the screen, or the source is implied to be present in the world, such as voices of characters and sounds made by objects in the world. It is an actual sound in the world that the characters could hear as well. Non-diegetic sound is coming from a source outside of the story space such as a narrator, score music or sound effects for dramatic effect. (Chion 1994, 73, 109). In videogames these lines are often blurred.

There are new terms to describe this such as trans-diegetic or informant diegetic (Jacobsen 2016), which we will elaborate more on later.

2.1.3 Evolving audio

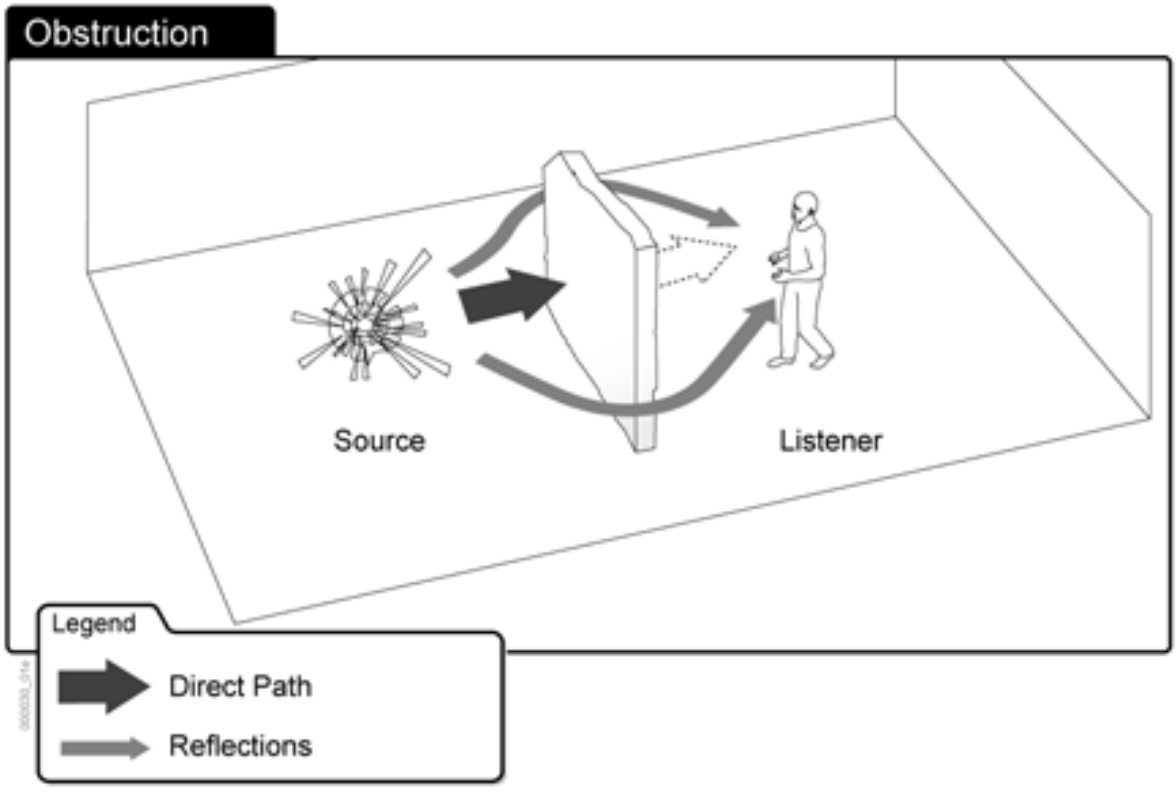
With the technical advancements made in the past few years, creating immersive, realistic, reactive and informant audio has become the new standard for video-games. This allows us to create more and more demanding audio systems that react to every action of the player in the world. (Neumann 2017). Special middle-ware software has been developed for creating more complex audio functionality such as Audiokinetic's Wwise or Firelight Technologies' FMOD. These help sound designers create complex audio systems with ease of an graphical user interface compared to hard-coding all of the functionality into the code (Collins, Kapralos & Tessler 2014).

Passing information through sound is no new convention but with the limitations of the first generation of videogame consoles, the amount of information might have been limited. In Super Mario Bros (Nintendo 1983) a simple uplifting ca-dence of note tells you what kind of a power-up came out of a brick when you jumped to break it. In Control (Remedy 2019) the Hiss is a prevalent part of the game and by sound alone it can tell you about what part of the game you're at, the state of the story and if there is an immediate threat. It is common for today's games to pass multiple layers of information through one sound, either directly influencing the player, or just enhancing the storytelling and immersion.

In modern games you might have a constant scanning of the environment and real-time analysis of the space around the player (Sorsa 2020). Modern systems for example create early reflections of the space in real-time. These are not only important in order to immerse the player in the world and the space that they are in, but they are important for spatial localization. The player can be backed up into a corner, but they necessarily don't have the time to look around where they can go next. Sound can be an important factor in informing the player of their surroundings.

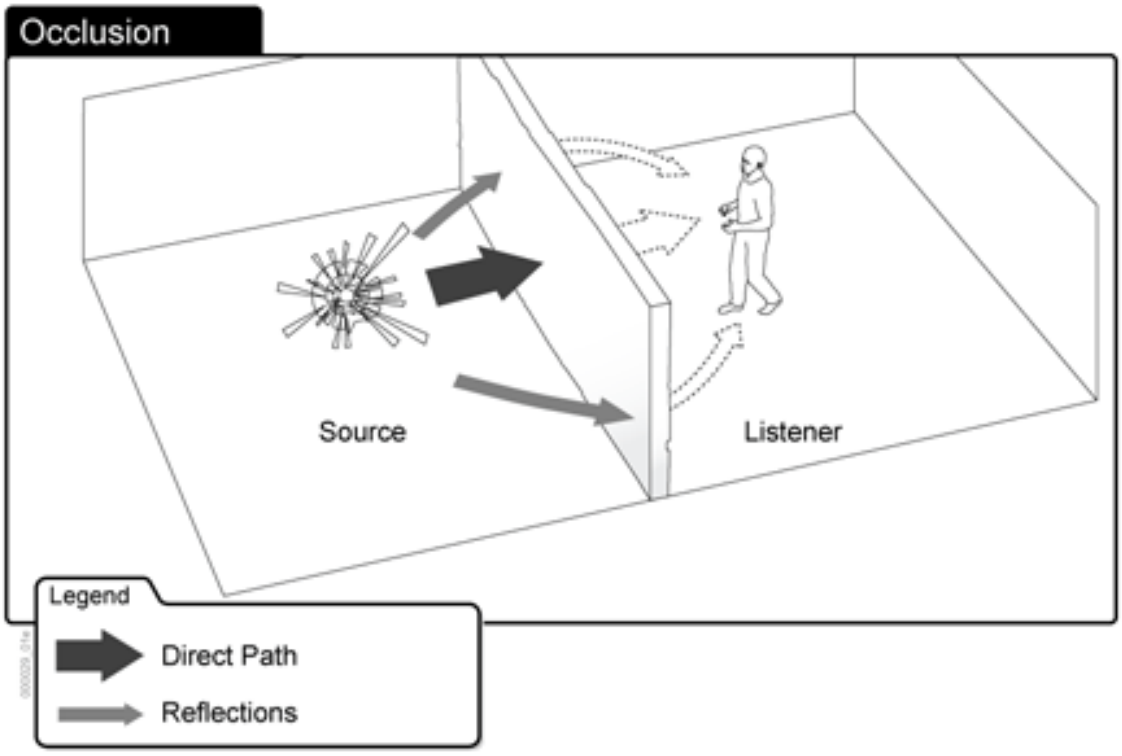
Occlusion, obstruction and propagation are also very important and calculated in real time. These all help the player to navigate and localize threats or object in the world. Propagating sounds correctly will help a player navigate through a maze by passing the audio through the correct pathway. Occlusion and obstruction likewise are equally important in guiding the player.

Obstruction means a partially blocked sound source. The direct path of the sound source would be affected but the indirect paths might remain the same, such as the reverberation of the room. In its most simplistic form this could be experienced as a muffled and quieter sound. In a game environment this could be reproduced simply with an on/off-switch that is based on direct line of sight to the sound source. When the player doesn't see the sound source you would have a low-pass-filter cut out some of the high-end frequencies and lower the volume a little bit. This is a very crude and basic system and does not necessarily work for every game. The higher the sense of immersion the more accurate all these parameters have to be. If the line of sight is blocked by a very narrow object that just happens to be right in between the listener and source, you wouldn't want to affect the sound as drastically. (Filion 2018, 280, 300).



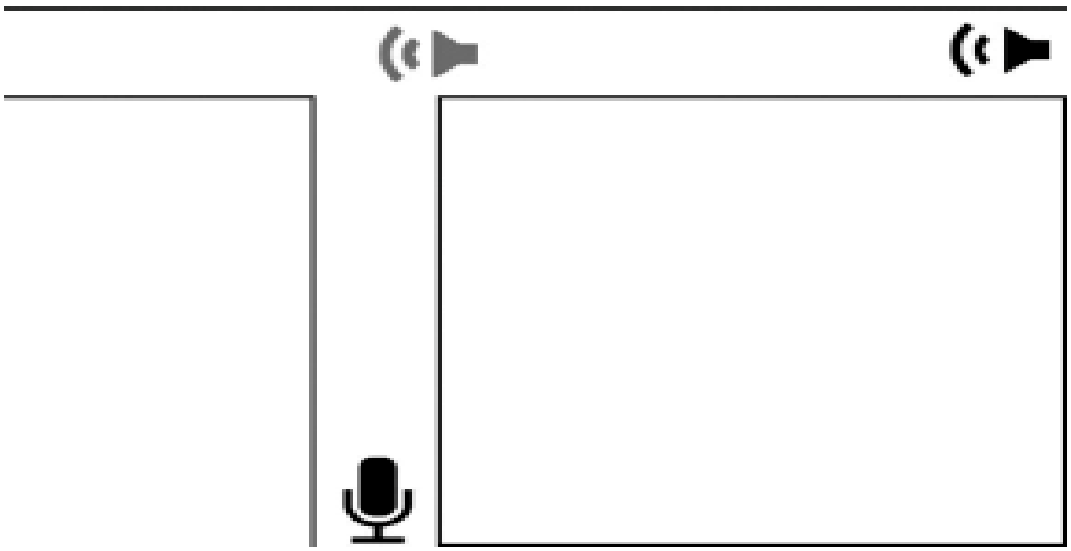
Picture 4. Example of Obstruction (Audiokinetic 2020).

Occlusion on the other hand means a completely blocked but still audible source (Audiokinetic 2020). This could for example someone talking in a different room behind a closed door, whereas if the door would be open, the sound would be merely obstructed. In an occluded sound you would also consider the reverberation of the space that the source is in and the resonant frequencies of the material the sound is passing through.



Picture 5. Example of Occlusion (Audiokinetic 2020).

More advanced systems would also check the material that's blocking the listener. Different materials pass sound through differently, so a concrete wall would muffle the sound a lot more than a wooden one. If you hear a sound through a wall like it would be straight in front of you, you may think the sound is originating from the wall. Instead you might want to detour the sound the to the closest opening, which leads us into propagation of sound. Propagation could be described as replicating the source sound in a closest possible direct path, with just a lower volume (Filion 2018, 282-283).



Picture 6. Example of sound propagation (Filion 2018).

The darker speaker to the right represents the original sound source and the microphone is the listener. You would not hear the sound coming a straight path through the wall but rather from the first possible opening at the end of the corridor as represented by the grey speaker. This is clear way to emphasize to the player that they must proceed to the end of the corridor first before trying to locate the sound source. With accurate propagation guiding the player using sound becomes much more efficient and realistic. (Filion 2018, 282-283).

3 GUIDING THE PLAYER IN AN OPEN ENVIRONMENT

3.1 Basics of level design

Good level design means that the player can navigate through the level in a designed fashion. The underlying rules of the level design should not be apparent to the player and the player shouldn't recognize patterns in the level design. In a 2013 GDC talk Dan Taylor from Square Enix Montreal, describes essential level design principals as follows:
"It should be fun to navigate, shouldn't rely on words, tell what but not how, constantly teaching, is surprising, empowers the player, has different difficulty levels, is efficient, creates emotion and is driven by mechanics. "

These are all integral qualities of good sound design, especially guiding and informative sound design. We want the player to naturally feel like they are coming up with the solutions and have natural reactions to the sounds in the game. An alarm should be an actual alarm in the game world and not a non-diegetic sound that they player recognizes as a designer's choice to warn you about something. It is a constant dialogue between the player and the game engine when they're exploring the level (Jacobsen 2016).

3.1.1 Traditional audio guiding techniques

In my interview with Ville Sorsa, lead sound designer at Remedy Entertainment, we talked about the guiding elements in their game Control (2019). Individual assets that guide the player to go investigate the source of the sound he called lure sounds. A traditional lure sound could be the sound of a hidden treasure so that the player understands that one is near and there is a puzzle involved in finding it. In Control there was also a character called Ahti that had strong musical leitmotif tied to his presence. This was also often used as a lure sound, Ahti's music would play in a specific space and the player would investigate knowing to expect something in there. (Sorsa, 2020).

In recent years there have been multiple titles released in the genre known as “Battle Royale”. The usual gist in these games is that the player is dropped into a restricted area without any or very little equipment and needs to gather resources quickly in order to survive. (Hornshaw 2019). In these types of games lure sounds can be used to help the player to locate the loot chests as efficiently as possible.

To emphasize an environments unique sound identity, an area could also be given a location ID, which means a sound unique to that environment (Sorsa 2020). In Control for example a reactor room could be defined by the sound of the reactor buzzing and humming. This helps not only with player immersion but with orientation. The player quickly recognizes by sound alone which room or space they are in. Rooms usually have their own reverberation set according to their size but having unique tonal SFX further emphasizes the difference between different spaces. (Sorsa 2020).

Voice-over, or VO, is also a part of passing information to the player (Sorsa 2020). The random chatter in a village could give the player an indication of the status quo in the world, without really having to change anything visually. The player character’s own VO is also used often as a guiding element. It could be used in an escalating manner such as *“I need to get higher”*, *“If only I could use something to get up there”, “I wonder if there is a ladder nearby”*. In Control, the by-standing NPC’s usually offer advice to the player character and not the player themselves, they also chat about the events that have unfolded due to the players actions.

In an interview for A Sound Effect, Mike Morasky from Valve was asked about the guiding elements in Valve’s latest release in 2020, Half-Life: Alyx. He mentioned struggling with keeping the player aware of the current objective without using too much dialogue. At a certain part of the game you are navigating towards a hotel. They gave the hotel a specific, repeating sound that could be heard from a long distance, so the player would be reminded of the current objective and the direction of the objective. “While you make the multi-hour journey through the hotel, you’re always aware of the sound of your goal and where it is in relation to where you’re going.” (Morasky 2020).

3.1.2 Using adaptive audio

The use of adaptive audio means that we can enhance the players reactivity with the world. The player can affect the soundscape of the world, but the world's sounds are also affecting the player (Collins 2013, 39). As the game progresses, we can pass information through audio alone even though the environment itself would be visually the same. Adaptive sounds enforce to the player that the things they do affect the world. This could be as simple as having a muffled state when the player is underwater or creating complex systems so that the music reacts to everything the player does. (Neumann 2017).

In Control, whenever you use a shield ability, the character Jesse, brings up a wall of rocks to protect the player. The surrounding environmental sounds are filtered so they appear more muffled to the player. This could be done just by using a simple low-pass filter but what's interesting here is that this adaptiveness affects the player in two ways. The calm, silent, filtered soundscape gives the player a sensation of not being in an immediate danger, kind of like when you run behind a pillar or a wall when an enemy is chasing you. Sound has been shown to have strong influence on the physical reaction of the player and that loud noises effectively can impair focus (Usher 2012). Silence on the other hand gives you a perception of space, in this example space for the player to collect their thoughts and think about their next move. Muffling the soundscape on the other hand means that the player cannot observe the environment as accurately using sound, as enemy footsteps and firing is muffled and visually impaired.

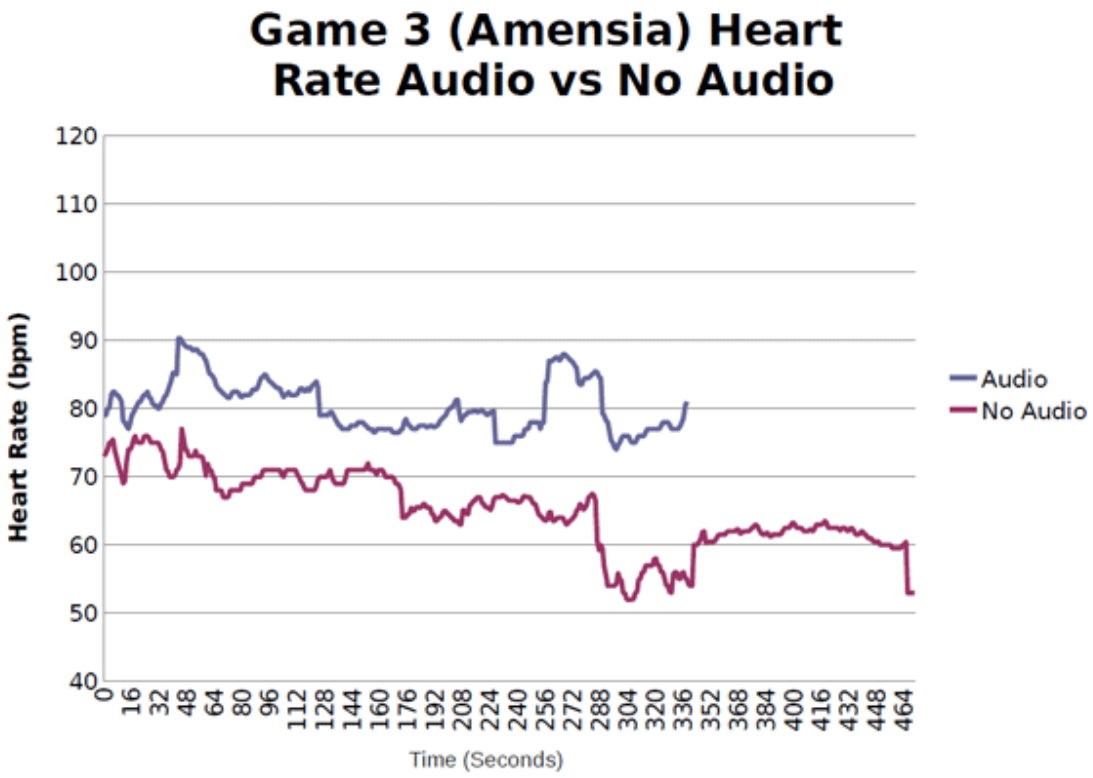


Figure 1. The comparison of heart rate with and without audio (Usher, R. Gamasutra 2012).

3.2 Diegetic sound

Diegetic sounds refer to sounds that are actual sounds (Chion 1994, 109). They are the sounds that the characters in the world space would also hear, such as footsteps, dialogue between character or music coming from a source, like a radio. Music that only the player hears such as background music is not diegetic. Diegetic sound is not only used to suspend disbelief, but it can also convey information by design, and this is what Bjorn Jacobsen would call Informant Diegetic in his thesis “Informant Diegesis in videogames”. The sound is not only needed for the actual events in the world, but they also can add to the information that the player receives or enhance storytelling.

3.2.1 Setting the tone

Besides just being a part of a scene, diegetic sounds also play a big part in setting the tone. Sound is a very effective way to tell something about an object that you cannot necessarily convey through visuals alone. For example, weight of an item

is hard to estimate visually, unless it is something that is commonly known as a heavy item. You could also emphasize the characteristics of an item by just its sound. A gun might have 100% accuracy in-game, meaning that it would always hit where you aim but if it sounds like it is falling apart when you shoot it, it affects the player experience. It affects how the player feels like the gun behaves and ultimately if and how they choose to use it (People Make Games 2019).

An example of this would be an online first-person shooter game called Wolfenstein: Enemy Territory. The game had two sides, Axis and Allies fighting each other. They both had their own unique models for their automatic rifles, Thompson for the Allies and MP40 for the Axis. The Lead Writer Ed Stern talked about why they had to make balance adjustments to the Thompson in an interview for a Youtube channel called People Make Games. Players felt that the Thompson was a lot more accurate and powerful than the MP40 but in reality, the guns had exactly the same behaviour and only differences were the visuals and the sound. The reason why people felt that the Thompson was a superior choice was because it had more low frequencies thus making it feel more powerful. (People Make Games 2019).

Using footsteps again as an example, the audio content of the footstep tells quite a lot about the character and the general setting in the scene. If we never even see the character that we are playing as, we can instantly tell by sound of the footstep what kind of a shoe we're possible wearing, how is the character built or what kind of clothes we are wearing.

3.2.2 Supporting artistic style

In Remedy's Control the environment is a transforming building called the Oldest House. Representing bleak bureaucratic buildings, the interior is concrete filled and brutalist. In my interview with Ville Sorsa he mentioned that they wanted to really emphasize this with the sound. As an example, he mentions that the reverbs are mixed 3dB louder than previous Remedy titles, just to really enhance the presence of the concrete. The game environment in Control seems like an

ordinary office building, but upon further inspection reveals to be seemingly endless, complex and surrealistically morphing. To enhance this combination of real and surreal, Sorsa talks about blending the diegetic and non-diegetic sounds together. This leads us into non-diegetic sound and how it blends with diegetic sounds in a videogame environment.

3.3 Non-Diegetic sound

Non-diegetic is something that has no real source in the world like score music, sound effects that emphasize drama or voice-over narration (Chion 1994, 73). It is something where the viewer is the receiving end of a monologue (Jacobsen 2015). It is a sound that is not in the world that the characters of the story are. Often in games you would have a special SFX for an important event. For example, in a shooter game where headshots do extra damage, a headshot could be emphasized by an additional 2D sound effect on top of the more diegetic bullet impact sound. This extra effect could be very stylized, iconic and limited in variation to help with recognition. This kind of extra layer would be mostly there to inform the player that they have indeed hit a headshot and is not necessarily heard by the character in the game, and would have a different attenuation, if any, compared to the more diegetic bullet impacts.

Dating all the way back to Super Mario Bros (Nintendo 1983) games have had different kinds of adaptive music systems (Collins 2008, 4). The character in the game does not hear the music, but a transition into combat music can nevertheless tell the player that there are enemies nearby. Sometimes, the characters even whistle the melodies in the background music which could be described as trans-diegetic. (Jorgensen 2007).

3.3.1 The power of implication

In the main quest of Control, there is a phone that seemingly has a source in the world, but it is also clearly audible through walls over unrealistic distances and stretches out into an unrecognizable drone when the player is further away. Es-

sentially the unrealistic distance on which the phone can be heard could be regarded as the character's intuition, a voice in their head. Once you get close enough to the phone it materializes into an actual source in the world thus changing from non-diegetic into diegetic. This is acting as a very obvious guiding element as the player constantly gets feedback that they are progressing in the game world. (Sorsa 2020).

In Control, there is the constant kind of non-diegetic presence of the Hiss, but during some parts of gameplay you can actually pin-point the source to loudspeakers in the game. In my interview with Sorsa we talked about this and he mentioned that it was actually a remnant from the original iteration of the Hiss, but it also tied the Hiss to the real world. It acted as a diegetic link between the absurd all-encompassing sound of the Hiss and the building to emphasize the story of the Hiss corrupting the whole building.

3.3.2 Passing information

Non-diegetic sounds are often used to emphasize a certain beat in the gameplay. Killing a last enemy might have an exaggerated sound effect or a long unrealistic reverb tail to let the player know that the combat is over. Discovering a secret may play a little musical stinger or a mysterious sound effect to emphasize that you have stumbled across something rare. It reinforces the experience and the player instinctively picks up that these are something that are hidden in the game and there are more to collect.

Music is also used a lot to pass information. Adaptive music system reacts to the events in the gameplay and you could have states for combat, stealth and exploring. These all pass information to the player either by design or even inadvertently, a lot of times the combat music might trigger before you even see any enemies. In Legend of Zelda: Ocarina of Time (Nintendo 1998) there is a forest segment where there is a very strong guiding musical element. In a maze-like sequence of the game, you have 4 different pathways to choose from. The only way to know for sure where each of them lead is through trial and error, but the music is directional to one of the paths and fades out if you head the wrong way.

It is a small detail, but it takes the guesswork out of the gameplay. It is also diegetic even though you can hear it over unrealistic distances. It is also a score theme in the game, but there is an actual source to the music, a character in the game is luring the player by playing it.

3.4 Examples

In this chapter a couple of game examples that have utilized sound as a guiding element will be studied further. One is a simple example of my own experience as a sound designer and my thoughts about the process, showcasing how a simple solution and design can affect gameplay. The other is Overwatch by Blizzard Entertainment, released in 2016, was developed with informative audio design in mind thus making it a prime candidate as an example (Lawlor 2016).

3.4.1 Overwatch

Overwatch by Blizzard Entertainment 2015 is a competitive team-based shooter. It has unique heroes each of them with unique abilities and it has 6 versus 6 objective based gameplay. In an environment like this, clear informative audio cues are imperative for gameplay, and Overwatch is one of the prime examples that was really designed with audio information in mind. Reportedly the Game Director Jeff Kaplan said early in development that he wanted to be able to play the game by sound, meaning that you could essentially play with no visuals at all. This makes the game a perfect case study for this thesis. In a keynote presentation for Audiokinetic in Montreal on June 20th 2016 at Centre Phi, Scott Lawlor, Project Audio Director and Tomas Neumann, Senior Software Engineer talked about the informative audio in Overwatch. The first audio design pillars for the game they defined were informative hero voice-over, a clear mix, pinpoint accuracy, gameplay information and Pavlovian response.

The voice-over, hereby referenced as VO, carries a lot of information. In Overwatch there are unique voices for each of the heroes but also an announcer, a kind of a meta-voice that is talking more directly to the player. The heroes also communicate in the diegetic sense to straight to each other, passing information about their status and objectives. This is emphasizing what is the objective or

right response to the current gameplay situation. They had created a VO priority system in order to enhance the information flow. Each of the characters have a lot of vocalizations, for example jumping and such, but you don't always want or need to hear these. The most basic limitation in the system is that not every sound is played to everyone. Things like the before-mentioned jumping is only audible to the player themselves. Very important sounds like Ultimate abilities VO and player death vocalizations are heard everywhere by everyone.

Having a clear mix is also an important feature in order to guide and inform the player in the most efficient way possible with sound (Lawlor 2016). In Overwatch the combat situations could get very intense and there could be 6 enemies shooting at you at once. This could get very overwhelming and turn into an audible mess very quickly. They created a set of rules that the engine is checking to determine a priority list for sounds. These were then divided into so called buckets that would only hold a certain amount of voices. For example, the highest bucket, the most immediate threat, could only have 1 character and their sounds in it. The game would ask questions like are they aiming at the player, how big is the players size on the enemy screen, for example a sniper aiming at them and are you shot at. These would all affect the priority number thus affecting the sound with various parameters to bring them in or out of player focus.

Lawlor and Neumann also talk about how setting up the occlusion and obstruction gave the players a sense of safety. Inside 4 walls they outside mix fades out and you can easily pinpoint if someone is nearby. Once you go out again the full mix comes back in and the flow of information is multiplied. Having very good spatial awareness with audio the players could also ignore enemies who couldn't hurt them. Just by audio alone you would know that the enemy is close, but you could also tell that they are for example downstairs. Like mentioned, very dangerous abilities such as ultimates ignored some of these rules in order to really emphasize the Pavlovian response here. The moment you hear the sound you have to react accordingly, so these sounds would play almost as 2D.

They also created a unique sound identity for each of the characters. Each character had their own set of foley, vocalizations, footsteps, abilities and attacks. By footsteps alone you could tell what hero was approaching. To emphasize the

Pavlovian response, Ultimate abilities would have very little variation by design in order to make them easy to recognize and learn to listen for those exact sounds.

3.4.2 Hero Siege

I worked as a sound designer on a project called Hero Siege by Panic Art Studios. It is a pixel graphic 2.5D ARPG in the style of Diablo II and essentially the core game loop is to kill enemies and get loot. It was common to have a screen full of enemies that you would want to wipe out as fast as possible, so there was a lot of information and a lot of it was either unnecessary or just missed. After a certain point of progress in the game, you no longer wanted to pick up nothing but the best items. This meant that for maximum efficiency you wanted to run through the map as quickly as possible while killing all the enemies, so under no circumstances did the player want to stop for anything unnecessary. This is why it was really important to make a unique sounds for the Satanic and Angelic drops, as they were the best drops in the game (Appendix 3).

Picture 7. Combat situation in Hero Siege. (Panic Art Studios 2019).

The sound design was approached from an emotional direction. The items had strong colourful aura tied to them when they dropped, red for Satanic and bright

yellow for Angelic. In order to make the sound cut through everything, the sound had to fill the whole spectrum. It needed to have a distinguishable attack and long enough of a body and decay to be heard over time as well. The drops were fairly rare in the game, so having only one variation for the sound was justified. It also emphasized the iconic nature of the sounds and player recognition. Back the game did not have separate volume sliders for separating different sounds in the game but eventually players were so fixated on only listening for the Satanic drop sound, that they had to be added to the game.

4 GAME PROJECTS

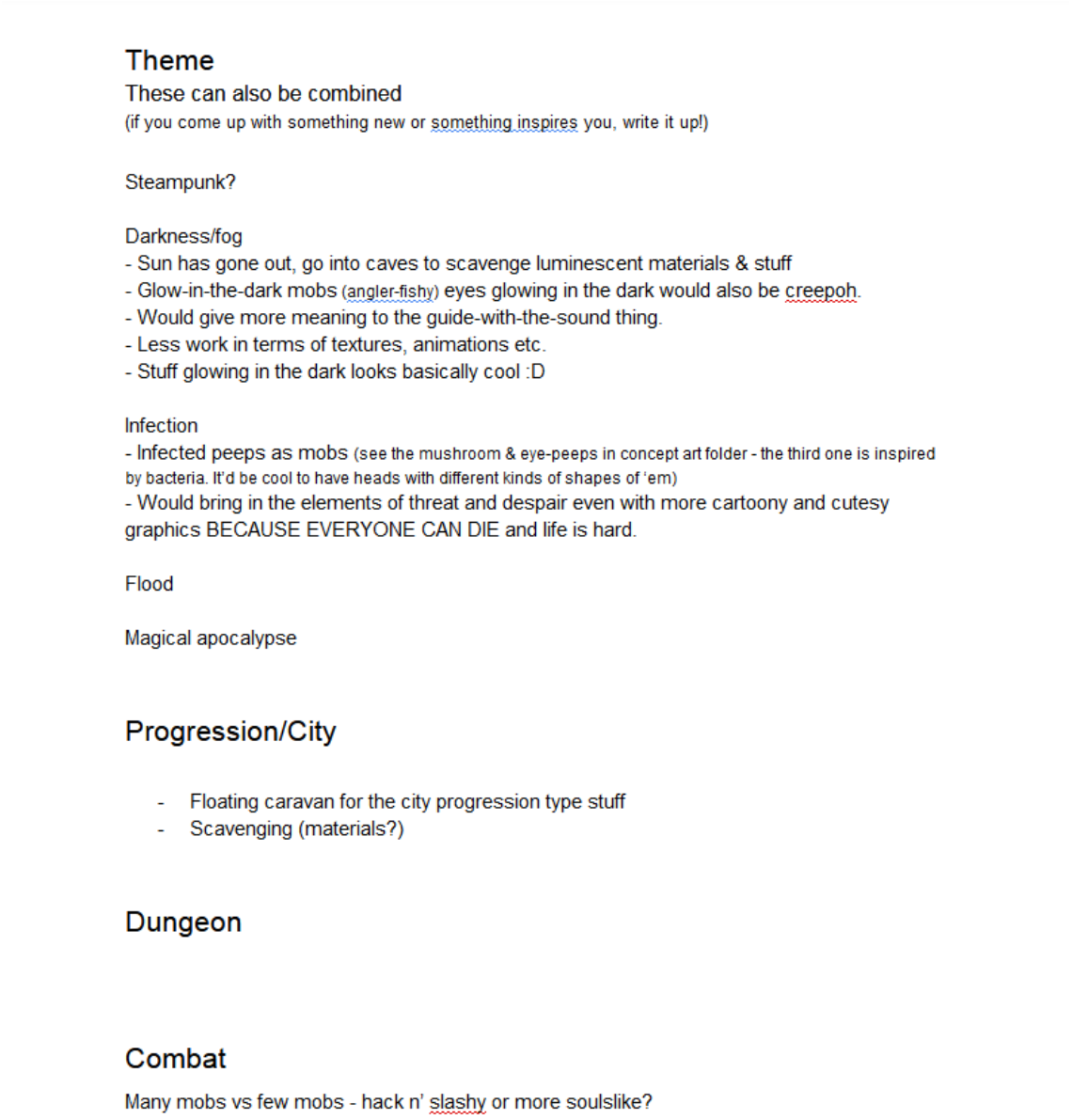
4.1 Project one: Dungeon Crawler

This practical part will demonstrate some real game world adaptations of the theories discussed. It will showcase the technical implementation and discuss the challenges that we as game sound designers face, compared to linear format. The team for the first project consists of 3 members, each of us doing our respective thesis on the same project. Justin Granger was our programmer, focusing his topic on procedural generation of terrains. Satu Kiintonen was doing graphics and focused on characters. Our main inspiration was a game called Dark Cloud, released in 2001 and developed by Level-5. The game combined aspects of rudimentary base-building and dungeon crawler elements. With these thesis topics in mind, we decided to focus on the dungeon crawling features of the game. The idea was to emphasize the role of audio by having limitations to player information. The world would be occluded by a heavy fog with a very minimal graphical user interface. This way a lot of the information would be conveyed through sound to further immerse the player in the world.

Due to the unpredictable nature of game development, this is also a reflection what can happen in the development process. By the time of writing the thesis the project may not be completely ready, but I will show my process as best as I can and discuss the remaining steps in theory.

4.1.1 Pre-production

Before opening any software or picking up any microphone, we must take a moment and think, what do we want to achieve with the sound. This always depends on the project and obviously in this case we wanted to demonstrate the leading capabilities of sound design.



Theme

These can also be combined

(if you come up with something new or something inspires you, write it up!)

Steampunk?

Darkness/fog
- Sun has gone out, go into caves to scavenge luminescent materials & stuff
- Glow-in-the-dark mobs (angler-fishy) eyes glowing in the dark would also be creepoh.
- Would give more meaning to the guide-with-the-sound thing.
- Less work in terms of textures, animations etc.
- Stuff glowing in the dark looks basically cool :D

Infection
- Infected peeps as mobs (see the mushroom & eye-peeps in concept art folder - the third one is inspired by bacteria. It'd be cool to have heads with different kinds of shapes of 'em)
- Would bring in the elements of threat and despair even with more cartoony and cutesy graphics BECAUSE EVERYONE CAN DIE and life is hard.

Flood

Magical apocalypse

Progression/City

- Floating caravan for the city progression type stuff
- Scavenging (materials?)

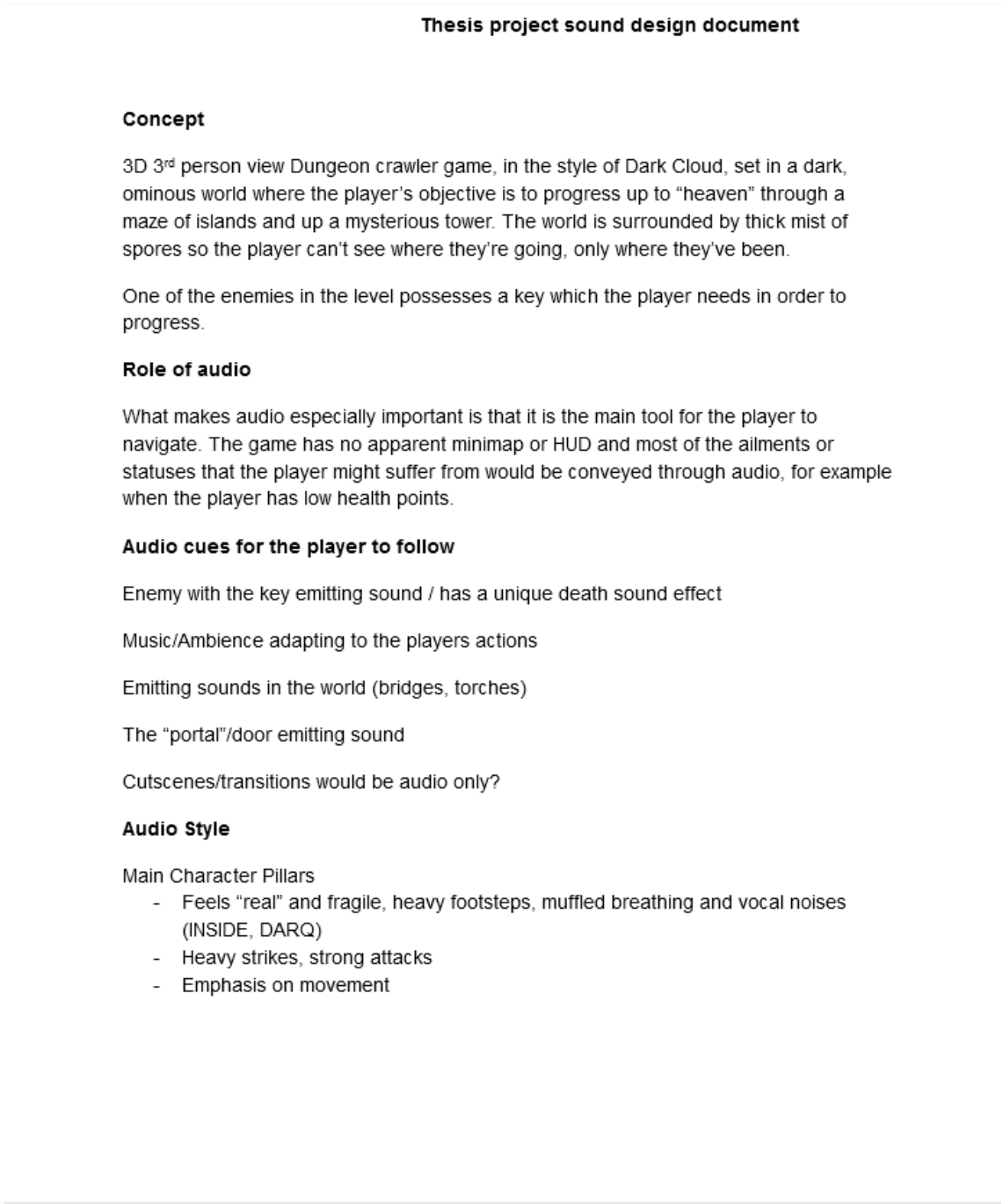
Dungeon

Combat

Many mobs vs few mobs - hack n' slashy or more soulslike?

Picture 8. Notes from one of our early meetings (Tolonen 2020).

After getting an idea of what we were doing with the game, I made a basic sound design document based on the previous meetings and the general consensus on the mood and atmosphere of the game. The idea of the document is to demonstrate overall idea of the soundscape, list the main pillars of the audio design and just have everything on a document as a basic guideline. This was also very important for establishing the main design principals for the game.



Thesis project sound design document

Concept

3D 3rd person view Dungeon crawler game, in the style of Dark Cloud, set in a dark, ominous world where the player's objective is to progress up to "heaven" through a maze of islands and up a mysterious tower. The world is surrounded by thick mist of spores so the player can't see where they're going, only where they've been.

One of the enemies in the level possesses a key which the player needs in order to progress.

Role of audio

What makes audio especially important is that it is the main tool for the player to navigate. The game has no apparent minimap or HUD and most of the ailments or statuses that the player might suffer from would be conveyed through audio, for example when the player has low health points.

Audio cues for the player to follow

Enemy with the key emitting sound / has a unique death sound effect

Music/Ambience adapting to the players actions

Emitting sounds in the world (bridges, torches)

The "portal"/door emitting sound

Cutscenes/transitions would be audio only?

Audio Style

Main Character Pillars

- Feels "real" and fragile, heavy footsteps, muffled breathing and vocal noises (INSIDE, DARQ)
- Heavy strikes, strong attacks
- Emphasis on movement

Picture 9. Sound design document draft (Tolonen 2020).

4.1.2 Design choices

After a few of meetings, we had decided on the general idea what the atmosphere was going to be like. Our graphics designer Satu then made some visual mood boards for us, based on what we had talked about. We had certain colour schemes and certain aesthetics where we were drawing influence from.

In an article called “Defining Your Creative Goals With Audio Storyboards” for Asoundeffect, Bryan Higa, Senior Sound Designer at Riot Games, talks about the value of creating an audio storyboard. He describes it as a good way to explore the fantasy and really try to dive into the story behind the sound identity. He mentions it is also a good way to get your development team on the same page about your vision and get them excited about what’s to come.

Next an audio storyboard mock-up was designed to try out the possible assets and the general soundscape of the world. In that storyboard I wanted to emphasize the barren world and the vicious combat the fragile player faces. This is also where the first drafts of the guiding elements in the game were demonstrated (Appendix 1).

The heavy cumbersome footsteps and the erratic, nervous breathing of the character are there to support the mood. Our hero is not immortal, and the armour has weight to it. He is afraid, lost and exhausted. You hear a distant howling, that is acting as a lure for the player as it is initially as a sound source, unknown to the player and fairly neutral by nature. Sudden devious laugh acts as a warning sign for the player. The creature suddenly attacks and kills the player. This sort of an audio mood board gave everyone something concrete to refer to when thinking about the soundscape.

Picture 10. Bryan Higa Audio Storyboard Example (Asoundeffect, 2019).

From there I started to flesh out the individual assets and to really define what are the guiding elements in the game.

4.1.3 Technical setup

The game was created using Unity Engine by Unity Technologies.
For this project I chose to use a middleware called FMOD from Firelight Technologies. This was so I could take of the audio as much as possible without having to make too many custom scripts or audio programming in Unity. This allows me to easily set up all the necessary functionality for the audio.

In FMOD I could easily control the attenuation curves, set up parameters to control various aspects of the audio functionality and have a lot of variations in the sounds themselves. For the implementation I would work together with our programmer to ensure proper functionality in the game engine.

4.1.4 Adaptive audio

This chapter demonstrates some of the sound features and how to implement and design these sounds into a real game project using FMOD and Unity. I decided to focus on a couple of key features and different types of audio guiding techniques.The game is heavily atmospheric and since I wasn't planning on doing music per se, it felt natural to have the ambience of the world to react to the players actions. Since the player was searching for something in a randomly generated world, it made sense to give subtle constant feedback to the player if they are going to the right direction. I wanted the ambience to subtly change depending on the players distance from the exit, but the feature couldn't be too noticeable to be actually distracting. The ambience would play everywhere in the map so the changes had to have more of a subconscious effect.

I drew inspiration for this from Quantum Break (Remedy, 2016). In Quantum Break they had a time machine that played the game's music either forward, if you're going in o the future or in reverse if you were going to the past. If you stood still in the time machine, the music would be timestretched, which is a common audio processing technique for making things seem "frozen".

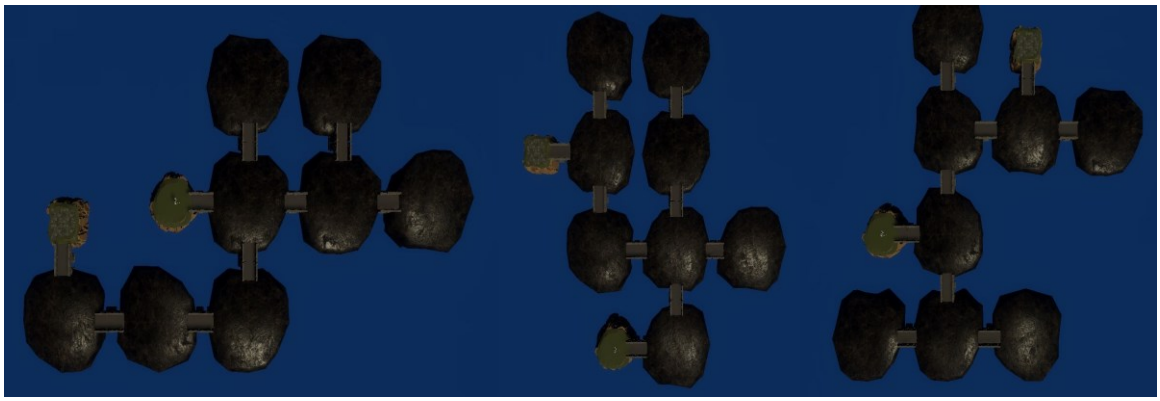
To further emphasize change in the state of the gameplay, when they player has found the key, and additional layer plays in the ambience. This way the player is reminded that their objective has changed and get information of the current state of the game, for example if they come back after a break, just by hearing the ambience you can tell where you left off. I also wanted the player to be able to locate the enemies easily and focus on them, so they are top priority in the playback hierarchy. Just to shift even more focus on the enemy sounds, they also duck other sounds.

4.1.5 Luring with sounds

The enemy carrying the key is the first clue the player is looking for. The sound is droned out over distance and slowly focuses on the enemy as the player gets

closer. This could be described as our first lure sound, it's something that the player is constantly trying to track and uses to progress through the level.

The environment consists of procedurally generated smaller islands, which are connected by bridges. In order to help the player to find the bridges on the edges of the island, these bridges also work as earcons or lure sounds. They all have emitters with a fairly short attenuation range, so you're not overwhelmed by bridge-sounds coming from all around you, but to just emphasize player attention to the importance of these bridges.



Picture 11. Top-down view of examples of randomly generated levels (Tolonen 2020).

Due to the procedurally generated nature of the level design, the emphasis is on different ways to help player navigation. In the picture above are some examples of possible randomly generated levels. The circular green area is the starting point and the player is trying to find their way to the square green area through these randomly generated connected islands.

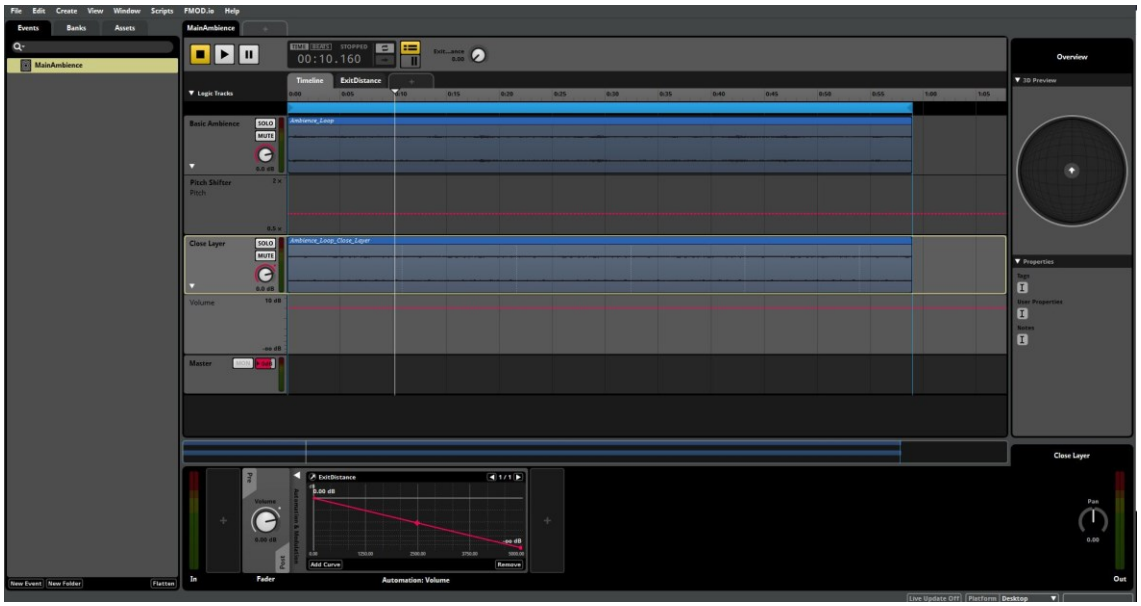
4.1.6 Implementation

Implementing the audio meant first designing the sounds themselves, designing their functionality in FMOD and finally inserting them into the game world in Unity. In FMOD you would determine the functionality of the sounds. For example, for the ambience that is changing based on the distance of player to the exit, I wanted to have a slight rising effect and an additional layer fading in the closer you are.

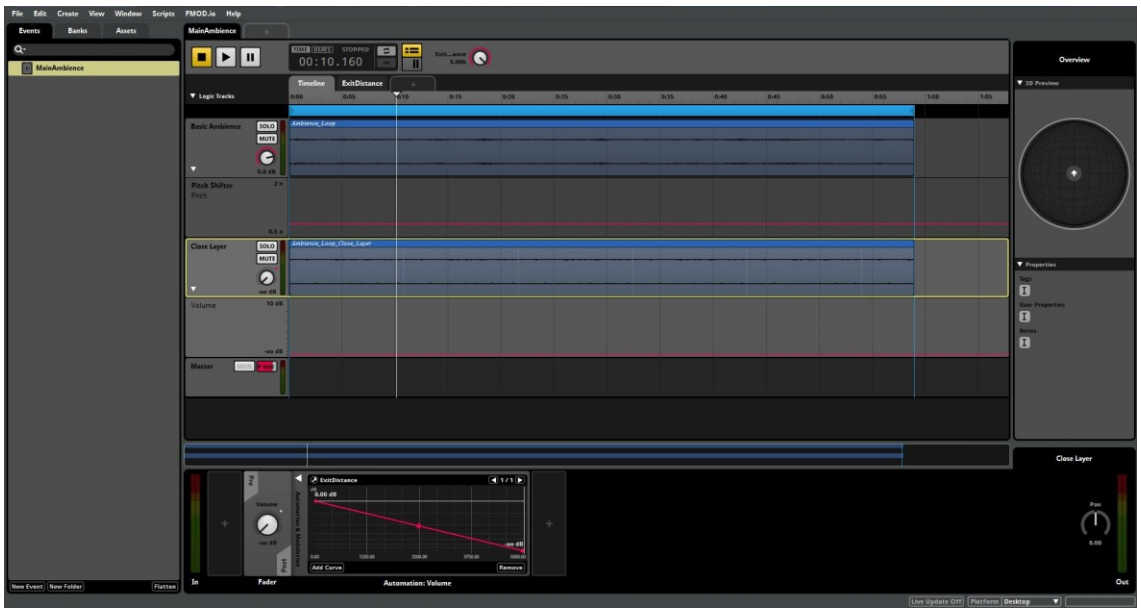
Implementation largely defines the functionality of the sounds. Things like attenuation and occlusion help with how to spatially localize the sounds and use them to guide the player (Lawlor 2016). In this project I'm mainly working with attenuation curves and effects that modify the sound.

4.1.7 FMOD & Unity

I created a basic 2D ambience loop that plays everywhere in the level. The parameter "ExitDistance" in FMOD is controlling the pitch of this initial layer from 0.8 to 1.0 value so the further you are the lower the pitch. The same parameter is also controlling the volume of an additional layer that is fading in the closer the player is to the exit. When the parameter is at 0, it means the player is the closest to the exit point. This is when the pitch is normal and the additional "Close-layer" is at full volume. The red lines in the pictures below represent pitch and volume automation respectively.



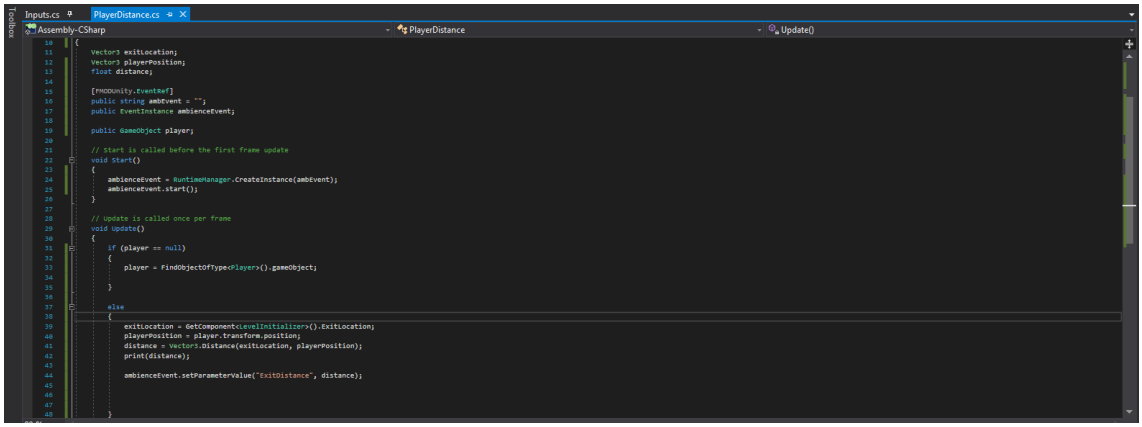
Picture 12. Ambience setup in layer at 0 distance (Tolonen 2020).



Picture 13. Ambience setup at maximum distance (Tolonen 2020).

The ambience is going to play everywhere in the map, so I don't have to worry about the attenuation of the sound. Usually you would determine attenuation for all of the sounds in the environment so they fall-off naturally in volume the further you are.

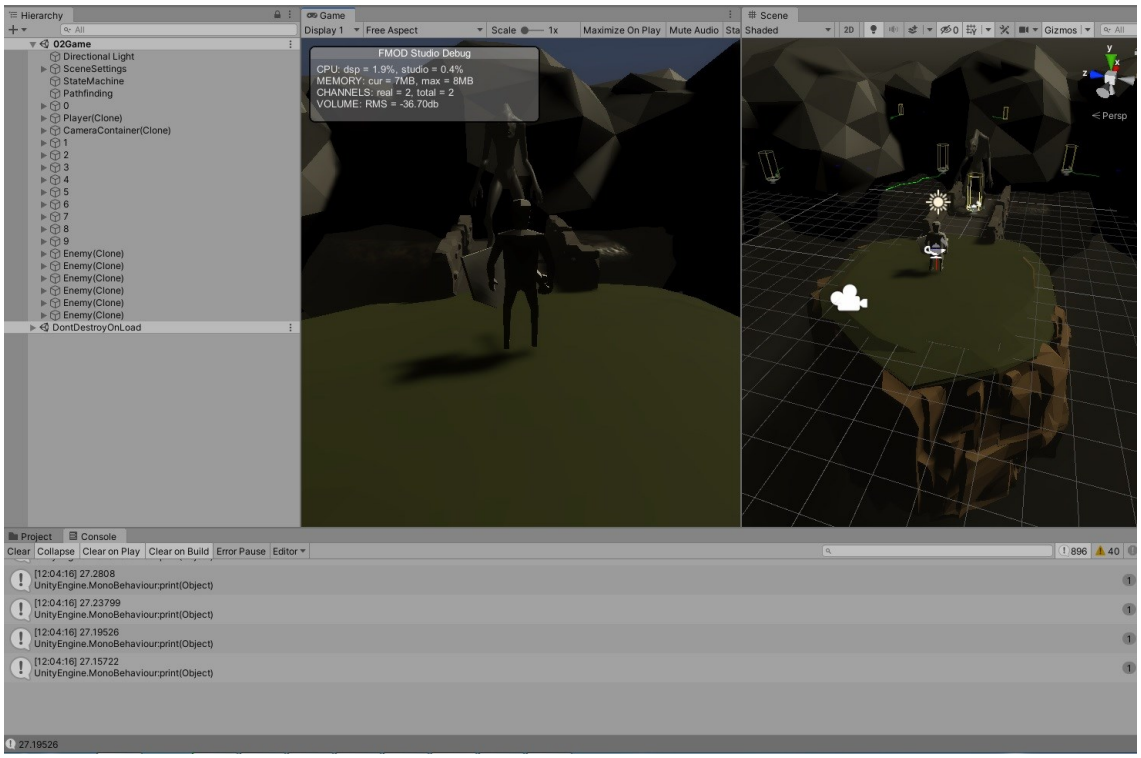
Next we need to determine how we are going to modify this sound. How will the player recognize the change? I wanted to use the player distance to the exit and emphasize the change in that. We can achieve this by calculating the distance between the two points and rounding that out to a usable value for our parameter in FMOD.



```
{
    Vector3 exitLocation;
    Vector3 playerPosition;
    float distance;

    [FMODUnity.EventRef]
    public string ambEvent = "";
    public EventInstance ambienceEvent;

    public GameObject player;

    // Start is called before the first frame update
    void Start()
    {
        ambienceEvent = RuntimeManager.CreateInstance(ambEvent);
        ambienceEvent.start();
    }

    // Update is called once per frame
    void Update()
    {
        if (player == null)
        {
            player = FindObjectOfType<Player>().gameObject;

        }

        else
        {
            exitLocation = GetComponent<LevelInitializer>().ExitLocation;
            playerPosition = player.transform.position;
            distance = Vector3.Distance(exitLocation, playerPosition);
            print(distance);

            ambienceEvent.setParameterValue("ExitDistance", distance);


        }
```

Picture 14. Code to calculate the distance (Tolonen 2020).

Without diving too much into the programming side of things, in the picture above we are essentially calculating the distance between the player and the exit on every frame. We would then use this value to drive the “ExitDistance” parameter in FMOD that was affecting our ambience SFX. After some testing I noticed the value is usually somewhere between 20-100, this meant I could set the maximum value of the “ExitDistance” parameter in FMOD to 100, as you would never be further away than that. See Appendix 4 for a video demonstration of the effect.



Picture 15. Using debug commands to print the distance value (Tolonen 2020).

4.2 Project two: Straight Shooter

The second project, project name “Straight Shooter”, is a personal custom game scene, showcasing a full soundscape and use of early and late reflections to guide the player. With the limited view of first-person, it might be challenging to locate where in the space the player is. In this game scene I am demonstrating using reflections from the gunfire to help the player understand their location in

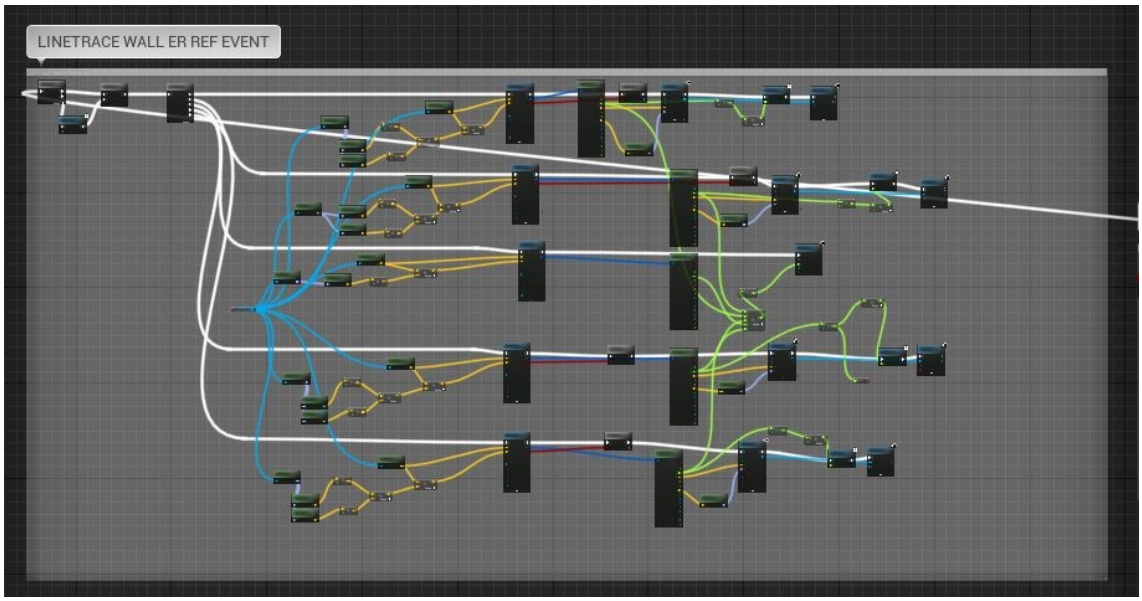
the space. The project was created with Unreal Engine 4 and Audiokinetic's Wwise, thus showcasing additional engine and middleware techniques. This project used 3D and gameplay assets from Unreal Engine Marketplace to create a realistic high-quality environment for audio demonstration purposes. All of the audio and its functionality was designed and implemented by me. The whole project includes informative features such as footstep sounds that change according to which material the player is walking on, grenade explosions that react to the environment around them and a dynamic reflection system for the gun. I will be providing a more in-depth look at the reflection system and how we can use that to inform and guide the player.

Picture 16. "Straight Shooter" (Tolonen 2020).

4.2.1 Reflection system

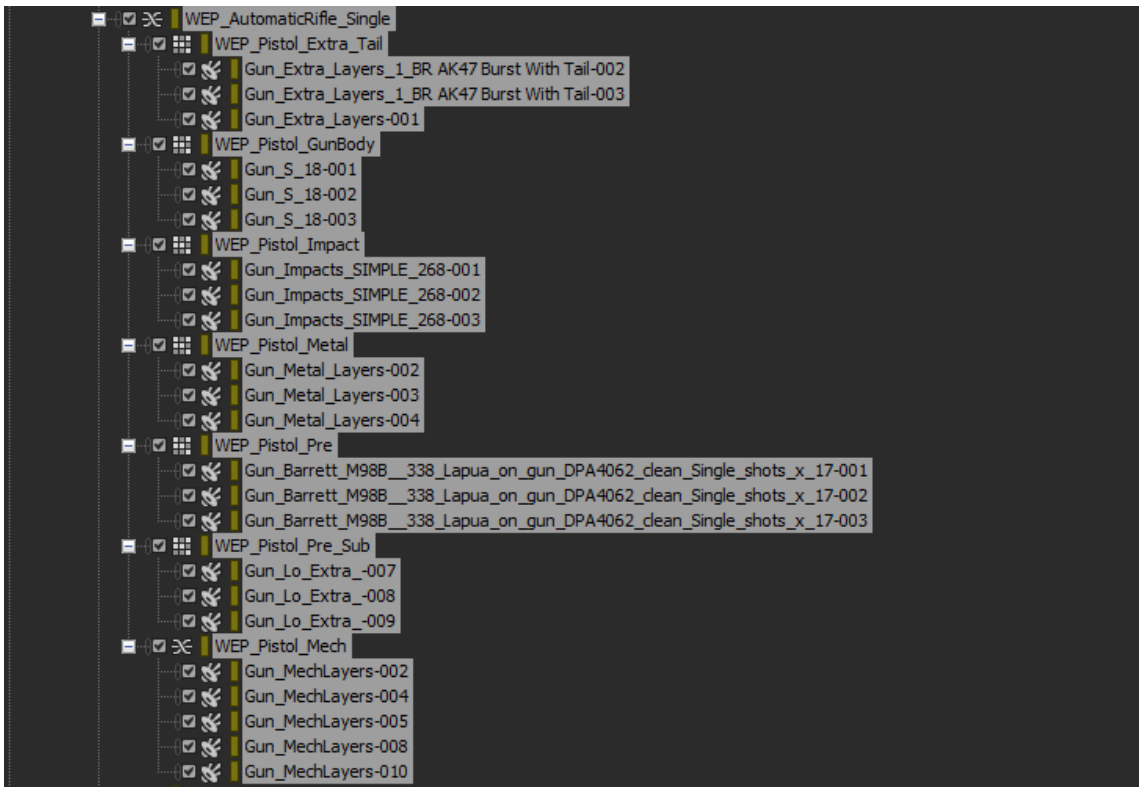
The goal was to create an immersive and dynamic reflection system for the gun firing. This is designed to immerse the player into the world by enhancing world reactivity with the player's actions so that the sound of the gun firing is reflecting off the urban environment. Reflections give the player a better understanding of the size of the space they are in and the distance to their nearest surroundings (McCaul 2017, 7). This is also designed to give importance and character to the gun firing, as it is powerful enough to generate energy to reflect off the environment. The gun is probably the most interesting and important sound in the game

scene so designing it to be pleasant but also informative is crucial. The actual implementation was created using Wwise and Blueprints in Unreal Engine. Blueprints are a node-based visual scripting system unique to Unreal Engine (Unreal Engine 2020).



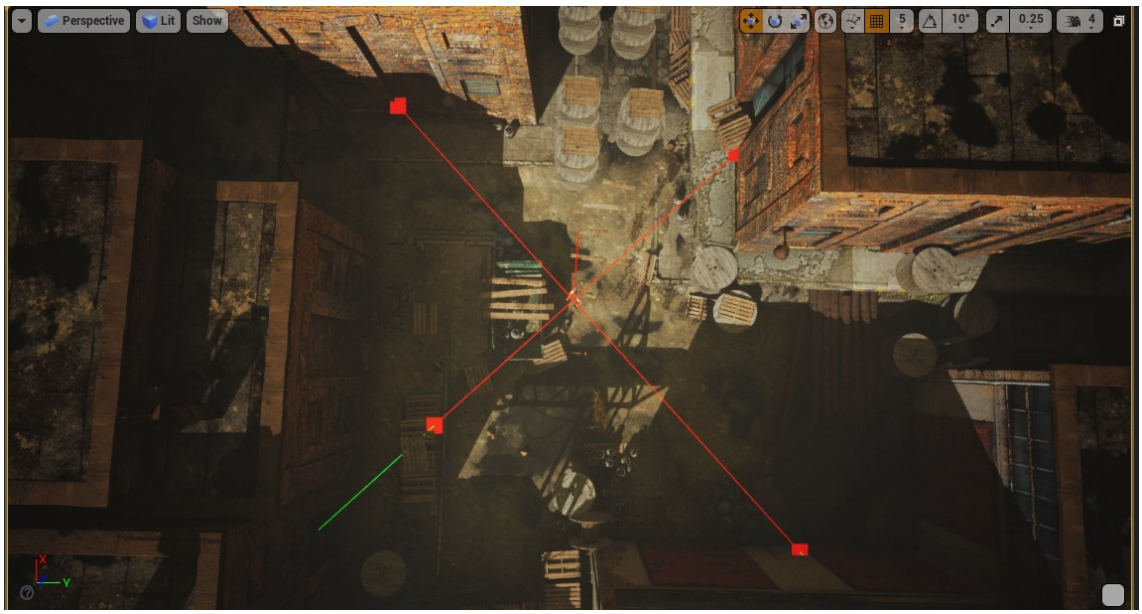
Picture 17. Blueprints of the reflection system (Tolonen 2020).

Using Wwise allowed me to create the behaviour I needed for the gun firing to function properly. Creating a single-shot sound with printed variations would get noticeable and jarring fairly quickly (Jacobsen 2018). Instead of having to pre-design dozens of different single-shot variations I broke down my design into layers, that each have multiple variations instead. This way I can organize them in a blend container, which is a container in Wwise than can play multiple layers of sound at once (Audiokinetic 2020). Now we have 7 layers that all play at once and each of them have several different variations inside of them. These variations are inside of a random container, that will then pick one of the assets inside of it by random. This way we have a lot of variation in the sound without having to pre-design all the variations.



Picture 18. Layers for the gun shot in Wwise hierarchy (Tolonen 2020).

For the reflections we are using the same assets as for the actual firing sound but with some layers removed, different attenuation and modulation, so that the player can differentiate between the original gunshot and the reflected sound. There is also a tail layer on top of the reflections for indoor and outdoor spaces that are there to emphasize the environment we are in. We are calculating distance from the player to the closest reflective surfaces in a quad-like formation around the player.



Picture 19. Top-down view above the player of the distance calculations with debug rays (Tolonen 2020).

In Picture 19 we can see the lines cast from the player. The boxes represent hits and reflections would play at each of those locations. The distance to the hit is calculated and the sound played is delayed by the distance divided by the speed of sound. Additional effects such as low-pass filtering and pitch are also applied to emphasize the change in the reflections according to the distance RTPC. We want the player to be able to recognize if the reflections are coming from near or far. We are also calculating if something is above the player to determine if we are completely encapsulated and thus changing the tail layer in the reflections into an indoor asset. See Appendix 5 for a demonstration of this change.

The reflections are telling the player how far or close they are to their surroundings (McCaul 2017, 7). In a first-person shooter like this the field of vision is limited and instead of forcing the player to constantly turn and see what is happening, we can use sound to inform the player of their surroundings. Even if you are running through the level with sound alone, you can tell if you are indoors or outdoors and how close to the nearest wall you are. In a tight combat situation this can be crucial for gameplay flow. Since the reflection sounds are delayed based on their distance, the player can spatialize the environment in a realistic manner (McCaul 2017, 7).

5 DISCUSSION

Sound is one of our earliest senses and I think it is fair to say that its importance cannot be belittled in any situation (Murch 2005). Sound gives context to everything happening around you in real life and in videogames. As shown, it can immerse the player into the gameworld, it can evoke emotions, it can offer information in a natural way and it can lead the player. It has the power to imply and tell a story beyond the visuals. You cannot see behind a corner, but you can hear it.

With more and more resources to spare in game development process and the hardware used, the attention to detail in sound is growing exponentially. We are constantly figuring out new and creative ways to enhance the player experience and the real-life guiding elements of sound are becoming more apparent in videogame environments as well (Morasky 2020). Game engines are also focusing more and more on creating better and more complex audio tools as awareness of audio's utility is on the rise. Gamers are getting better and better equipment in their homes and are enjoying the medium in a very high-fidelity environment and they understand the level of quality for good audio.

In my example game projects I think I have accomplished proving that sound can be and is a valuable tool in any game development project. If we are limiting the information flow to the player through non-diegetic instruments such as graphical user interfaces, tooltips or tutorials we are creating a more immersive world but at the same time we need to find a natural alternative way to convey information. This is where sound excels and it does not break the immersion, it is all part of the world. Unfortunately, due to the challenges faced in the game development process, we didn't get as far with the game as we would have hoped to. No playable build has been published at of the date of writing but we hope to continue the project to a published work. In the second project, Straight Shooter, the effect of guiding is more subtle and less of an artistic choice but it is a technically more advanced example. It provides the player with valuable spatial information without having to use additional graphical elements. Combining both creative and technical solutions is important in order to create a cohesive and functional, informative soundscape for a game of any calibre.

REFERENCES

Audiokinetic .2020.Wwise Documentation. Read on 12.5.2020 https://www.audiokinetic.com/library/edge/?source=Help&id=welcome_to_wwise

Adams, E. 2006. Fundamentals of Game Design. New Riders Publishing.

Chikhani, R. The History of Gaming: An Evolving Community. Techcrunch, 2015. Read on 16.4.2020 https://techcrunch.com/2015/10/31/the-history-of-gaming-an-evolving-community/

Chion, M. 1994. Audio-vision: sound on screen. New York: Columbia University Press.

Collins, K. 2008. Game Sound: An Introduction to the History, Theory and Practice of Video Game Music and Sound Design. Cambridge, Massachusetts: The MIT Press.

Collins, K. Game Sound in the Mechanical Arcades: An Audio Archaeology. Game Studies, Volume 16, issue 1, October 2016. Read on 12.2.2020 http://gamestudies.org/1601/articles/collins

Collins, K. 2013. Playing with sound: a theory of interacting with sound and music in video games. Cambridge, Massachusetts: The MIT Press.

Collins, K. Kapralos, B. Tessler, H. 1994. The Oxford Handbook of Interactive Audio. Oxford University Press.

Filion, M. 2018. Obstruction, Occlusion and Propagation. In Somberg, G. (ed.) Game Audio Programming 2: Principles and Practices. CRC Press.

Füsslin, F. 2017. Working with Audio Designers. In Somberg, G. (ed.) Game Audio Programming: Principles and Practices. CRC Press.

Graetz, J. 1981. The Origin of Spacewar. Read on 11.2.2020. http://www.wheels.org/spacewar/creative/SpacewarOrigin.html

Grimshaw, M. 2008. Sound and Immersion in the First-Person Shooter: Mixed Measurement of the Player's Sonic Experience.

Higa, B. Defining your creative goals with audio storyboards. A Sound Effect, 2019. Read on 20.3.2020 https://www.asoundeffect.com/defining-your-creative-goals-with-audio-storyboards/

Hornshaw, P. The history of Battle Royale: From mod to worldwide phenomenon. Digital Trends, 2019. Read on 11.3.2020. https://www.digitaltrends.com/gaming/history-of-battle-royale-games/

Jacobsen, B. 2015. Video game and real life audio design as game design. Presentation. SPILBAR 23. Viewed on 3.3.2020
https://www.youtube.com/watch?v=d1yH92UYptw&

Jacobsen, B. Game Audio Immersion. A Sound Effect, 2018. Read on 5.3.2020
https://www.asoundeffect.com/game-audio-immersion/

Jacobsen, B. 2016. Informant Diegesis in Videogames. Aarhus University.

Jorgensen, K. 2007. On Transdiegetic Sounds in Computer Games. University of Bergen.

Jorgensen, K. 2008. In Collins, K. From Pac-Man to Pop Music: Interactive Audio in Games and New Media.

Lawlor,S & Neumann, T. 2016. Wwise Tour 2016 – Blizzard Overwatch (1-7) – Audio in Overwatch. Key-note presentation. Viewed on 4.3.2020
https://www.youtube.com/watch?v=fGDOgn2WXLs

McCaul, S. 2017. Principles of Audio. In Somberg, G. (ed.) Game Audio Programming : Principles and Practices. CRC Press.

Morasky, M. Composer. 2020. Designing Half-Life: Alyx's Superb VR Sound. Read on 18.4.2020.
https://www.asoundeffect.com/half-life-alyx-vr-sound/

Murch, W. 2005. Dense Clarity – Clear Density. Read on 3.2.2020.
https://transom.org/2005/walter-murch/

Neumann, T. 2017. Dynamic Game Data. In Somberg, G. (ed.) Game Audio Programming: Principles and Practices. CRC Press.

Unreal Engine Documentation. Blueprints Visual Scripting. Read on 12.5.2020.
https://docs.unrealengine.com/en-US/Engine/Blueprints/index.html

Sorsa V. Lead Audio Designer at Remedy Entertainment, Interview on 20.2.2019, Interviewer Tolonen, J. Transcribed. Tampere University of Applied Sciences, Mediapolis campus.

Taylor, D. 2013. Ten Principles for Good Level Design. Key-note presentation. GDC, San Francisco. Viewed 10.2.2020
https://www.youtube.com/watch?v=iNEe3KhMvXM

Usher, R. How Does In-Game Audio Affect Players? Gamasutra, 2012. Read on 14.2.2020
https://www.gamasutra.com/view/feature/168731/how_does_ingame_audio_affect_.php

Why the sound of a gun had to be nerfed in Wolfenstein: Enemy Territory. People Make Games. Youtube video. Published on 29.8.2019. Viewed on 16.2.2020 https://www.youtube.com/watch?v=RDxiuHdR_T4

GAMES

Control. 2019. Remedy Entertainment.

Dark Cloud. 2001. Level-5.

Diablo II. 2001. Blizzard Entertainment.

Hero Siege. 2014. Panic Art Studios.

Overwatch. 2016. Blizzard Entertainment.

Pong. 1972. Atari

Quantum Break. 2016. Remedy Entertainment.

Spacewar! 1962. Steve Russell

Super Mario Bros. 1983. Nintendo.

The Legend of Zelda: Ocarina of Time. 1998. Nintendo

Wolfenstein: Enemy Territory. 2003. Splash Damage.

APPENDICES

Appendix 1. Audio Storyboard Demo

https://soundcloud.com/juuso-tolonen/audio-storyboard-demo/s-eeb9UDb8IHQ

Appendix 2. Interview with Ville Sorsa

https://drive.google.com/open?id=1uQhK8-cONrbnwjKo1P2y6RmQZPNS_1O_

Appendix 3. Hero Siege Satanic Drop SFX

https://www.youtube.com/watch?v=CMqKGHeWE0o

Appendix 4. Demonstration of distance-based ambience modulation.

https://www.youtube.com/watch?v=bkVMf8--V-c

Appendix 5. Gameplay footage of Straight Shooter

https://www.youtube.com/watch?v=3PQpCQLnx1Q